"Controlling Beauty"

Perceptual Representation and Control of Geometric Shape Spaces

P. Pundir Y. Wu S. Soatto

‡ Department of Computer Science, University of California, Los Angeles – CA 90095
† Department of Electrical Engineering, Washington University, St.Louis – MO 63130
Tel: (310)825-4840, Fax: (310)794-5057, email soatto@ucla.edu

Abstract

We consider the problem of representing geometric entities in a perceptual coordinate system. A geometric object (for instance a closed contour or a surface patch) described by a finite number of parameters is mapped onto an arbitrary collection of perceptual coordinates via a psychophysical experiment, resulting in a collection of marginal probability distributions. Viceversa, a set of perceptual coordinates can be mapped onto a probability distribution of parameters, and a pointwise inverse can be constructed by tracing the mode(s) of the density. We use such a map to modify the shape of an object by controlling perceptual characteristics, as opposed to manipulating the geometry directly. We choose to test our hypothesis on a simple one-dimensional experiment to map the shape of eyeglass frames to the perceived character of the person wearing them. We believe that perception-driven geometric manipulation presents vast opportunities in a number of commercial applications of mass-customized design.

1 Introduction

Physical objects elicit perceptual responses. For instance, cars differ in their appearance as one may look more "aggressive" or less "goofy" or more "sporty" (Figure 1). Similarly, a face – intended as a particular surface in space supporting a particular radiance distribution function – may appear attractive, intelligent, driven, intense etc. depending upon its particular geometry and photometry. Even for the same face, subtle changes such as the shape of eyeglasses can alter the perceived character (Figure 1). Such responses may be quantified subjectively in terms of the "amount" of certain perceptual characteristics, or "labels", along with a choice of labels themselves. Since the perceptual response is subjective, a given object presented to a number of human subjects elicits *not* an individual response, but a (marginal) *distribution* on each label, as suggested in Figure 2. Naturally, such a distribution depends upon a number of nuisances due to cultural, historical, geographical factors, and it is not time-invariant: a car that looked aggressive to a thirty five year-old lawyer in Spain in the sixties my appear tame to a twenty year-old asian actress in California today, although it may still look aggressive to a teen-aged male student in Nepal. Overlooking nuisance factors can be dangerous, for it can lead one to believe that there is a one-to-one "universal" map between geometric and perceptual features. This point of view was taken to an extreme by the pseudo-scientific movement of Phrenology in the nineteenth century [**?**, **?**]. The proponents of Phrenology did not limit themselves to believe that there is a map between geometric features and perceived character; they believed that there is a map between physiognomic features and personality traits. For instance, if the slope of somebody's forehead was within certain values the person was a potential murderer!

In this paper we take a far less dogmatic approach: we consider different representations of a geometric shape space (for instance the space of closed piecewise smooth curves after factoring out the Euclidean group) and use data to derive maps between distributions in different representations. Although the methodology we describe is general, in the interest of clarity we illustrate our ideas on a one-dimensional example, that is the design of eyeglass frames based on perceptual, as opposed to geometric, specifications.

1.1 Applications

Giving a perceptual coordinatization of a geometric space – characterizing the map between geometry and perception – can be useful in a wide variety of applications. For instance, how can a plastic surgeon shape the nose of a patient so that she can

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE **2001**	2. REPORT TYPE	3. DATES COVERED **00-00-2001 to 00-00-2001**
4. TITLE AND SUBTITLE **'Controlling Beauty' Perceptual Representation and Control of Geometric Shape Spaces**		5a. CONTRACT NUMBER 5b. GRANT NUMBER 5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)		5d. PROJECT NUMBER 5e. TASK NUMBER 5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **University of California, Los Angeles,Computer Science Department,Los Angeles,CA,90095**		8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)		10. SPONSOR/MONITOR'S ACRONYM(S) 11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT
see report

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**Same as Report (SAR)**	**10**	

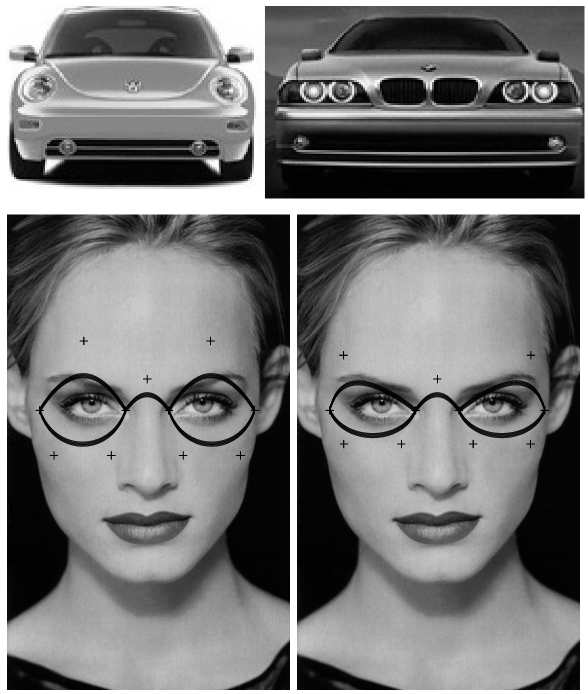

Figure 1: *Friendly? aggressive?*

look more assertive? How does a makeup artist make an actor look more dramatic? How can one modify clothes so that, once worn, they make the person look taller, or - lest we say - less fat? In all of these cases, one would like to determine how the geometry needs to be modified in order to achieve a given perceptual goal. This can be extended to manufacts: for instance, in designing eyeglass frames, how does the shape need to change in order for the final product to have a given perceptual characteristic? How does an industrial designer modify the tail-lights of an automobile to make it appear more sporty or classy? Such a knowledge is currently domain of artists and designers, who exhert it through a process that seems to defy formalization. Certain labels (for instance "beauty") that are very broad and losely defined will have a very broad geometric correlate, and will therefore correspond to broad densities. However, we believe that in well-defined and restricted domains it can be made analytical and, to a certain extent, be controlled.

In addition to mapping static shapes to perceptual characteristics, one may want to map *dynamic* models. For instance, how can one adjust the dynamics of a moving articulated body (e.g. a simulated human character) so as to make it appear assertive, or tentative, or drunk?.

Our goal is to be able to place all these questions on a solid quantitative footing. Doing so requires a blend of geometry, statistics, dynamical systems and perception psychology. While most of the data will be gathered in controlled experiments, artistic knowledge and experience can still be used to guide the experiments and the data analysis.

1.2 Contributions of this paper and relation to previous work

This paper poses the problem of representing a finite-dimensional shape using a finite dimensional set of perceptual labels. This problem is present "in nuce" in several statistical techniques, for instance collaborative filtering [**?**], although to the best of our knowledge the attempt to build an explicit map between the two representations is novel.

Of course this paper relates to the broad literature on perception and representation; although we do not have space to venture into an extensive review, for a historical perspective the reader can refer to [**?**, **?**]. On the relation between design and

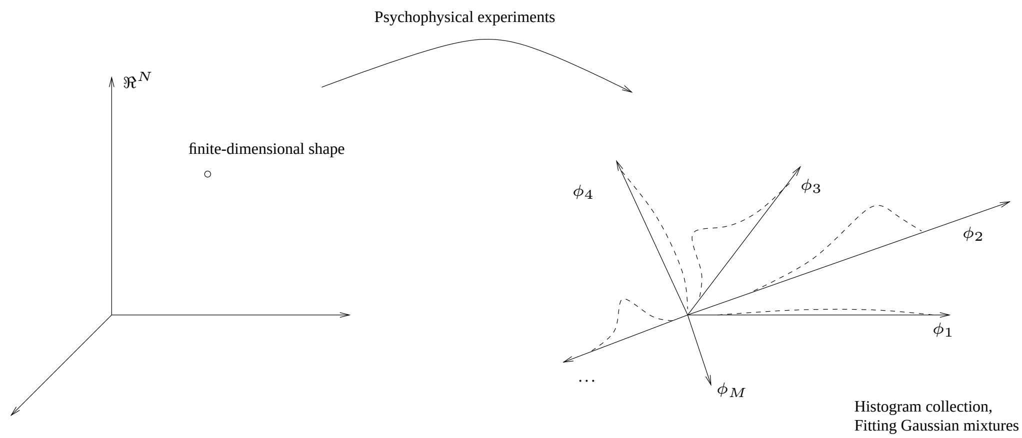

Figure 2: *A shape described by a finite number of parameters may be represented as a point in a linear vector space (left). When presented to a number of human subjects, it induces a distribution of perceptual characteristics or "labels".*

perception one can see for instance [**?**]. This work is losely related to work in color vision where one is interested in mapping color names ("labels") to either physical properties or neural responses, and to work in perceptual categories that we do not review in detail here.

A research agenda similar to ours has been carried out in [**?**] for the case of human faces, where a test image is projected onto the span of a set of solid models acquired with 3D scanners and labeled by human observers. The "character" of the model can then be modified by increasing the weight of faces labeled with that character in a linear combination.

While the "direct" map from geometric to perceptual coordinates is built by means of a straightforward psychophysical experiments, the inverse is a set map and is not easily characterized. In a probabilistic sense, the experiment can be characterized by a joint probability density between geometric features and labels; given a prior on labels, the experiment determines the posterior marginal densities, and inversion can be performed as usual using Bayes' rule. However, since we are interested in a point-inverse, we propose deriving a reduced map by considering the maximum of the probability density defined by the inverse map when such a maximum exists and is unique. When the pre-image density is multi-modal, we do not define an inverse map. How to choose independent labels in such a way that their pre-image densities are unimodal is an interesting question that we leave for future research.

We test our hypothesis on a simple one-dimensional example that relates to the shape of eyeglasses and their impact on the perceived character of the person wearing them. We characterize the shape of the eyeglasses as a closed contour and allow only one degree of freedom.

2 Perceptual representation of geometric features

Consider a geometric object, such as a collection of curves and surfaces, represented by a finite number N of parameters $x \in G \subset \mathbb{R}^N$. These can be the parameters of a CAD model for a product, or the parametric description of a set of geometric features. Now define a number M of "labels" as perceptual characteristics $\phi_1, \dots, \phi_M$. These are functions

$$\phi_i : \mathbb{R}^N \longrightarrow [0,\ 1];\ x \mapsto \phi_i(x)$$

onto the normalized "intensity" of label i. Such labels are part of the design process and depend upon the domain of application. For instance, a car manufacturer may choose the adjectives that are most expected in a particular brand or market segment. Note that the labels are not necessarily independent, in the sense of elements in a vector space: given a point ψ there may be different combinations of labels that represent it:

$$\psi = \sum_{i=1}^{M} \alpha_i \phi_i = \sum_{i=1}^{M} \beta_i \phi_i$$

where $\alpha_i \neq \beta_j$ for at least one pair (i, j). Consider now *one* experiment, where for a particular point $x \in G$ a subject is asked – in a forced choice paradigm – to assign a value for each label. This experiment can be represented as a map from G to the M-unit cube $[0,\ 1]^M$:

$$\mathcal{E} : G \longrightarrow [0,\ 1]^M;\ x \mapsto (\phi_1(x), \dots, \phi_M(x)) \doteq y.$$

Naturally, the experiment is not an injective map. Indeed, the pre-image of the experiment $\mathcal{E}^{-1}(y) \subset G$ can be any subset of G, not necessarily connected. For instance, many different faces may be perceived as being equally assertive even though their physiognomy is very different.

3 Mapping distributions

Consider now not one but K experiments $\mathcal{E}_1, \dots, \mathcal{E}_K$ all mapping the same point $x \in G$ onto the unit cube Φ. The ensemble of experiments, $\mathcal{E}^K$ maps G onto ensembles of values in Φ. Such ensembles can be organized into one histogram h_i for each label ϕ_i:

$$\mathcal{E}^K : G \longrightarrow \mathcal{P}^M([0,\ 1]);\ x \mapsto (h_1(x), \dots, h_M(x))$$

where $h_i : [0,\ 1] \rightarrow [0,\ K]$. The normalized histograms can be approximated by probability density functions $p_i,\ i = 1 \dots M$ defined on the unit interval, for instance using Gaussian mixtures. Consequently, an ensemble experiment for a particular shape $x \in G$ can be represented as a map from G to the set of marginal densities p_i. Some of these densities may be uninformative (i.e. uniform in the unit interval), indicating that the particular point x does not elicit a response on label i. Some may be multi-modal, and yet others may be concentrated around a single mode. In general, one can expect the joint density functions to be multi-modal, uninformative along certain directions, and highly complex.

Consider now performing the repeated experiment $\mathcal{E}^K$ on each point of a discrete lattice on G. Let $x_i,\ i = 1 \dots l$ be the points on the lattice. Correspondingly, after performing the experiment we collect a set of densities $p_i^j,\ i = 1 \dots M, j = 1 \dots l$ that summarizes the experiment as a map from G to Φ.

From a statistical point of view, the underlying assumption is that there is a joint probability density $p(x, \phi)$ and that, given a prior on the coordinates x, the experiment is described by the density $p(\phi|x)$, of which we measure the marginals. The inverse in probabilistic terms is just $p(x|\phi)$ which can be computed using Bayes' rule. We are insterested in using this model to extract a point-map for the inverse. To this end we consider the map that associates a certain set of coordinates on Φ (not just a particular point on Φ, since it may have many different coordinatizations) to points on the lattice that correspond to it via the experiment:

$$(\mathcal{E}^K)^{-1} : [0,\ 1]^K \longrightarrow G;\ y \mapsto \{x \mid (\phi_1(x), \dots, \phi_M(x)) = y\}.$$

From these pre-images, we construct a histogram in G (by evaluating the density of points in the pre-image) and fit a density using a Gaussian mixture. This construction defines a map from points on Φ onto densities on G:

$$(\mathcal{E}^K)^{-1} : \Phi :\longrightarrow \mathcal{P}(G);\ y \mapsto g \text{ where } g : G \rightarrow [0,\ 1].$$

This is illustrated pictorially in Figure 3. When a point moves in perceptual space, the mode(s) of the induced density will describe a trajectory in geometric space. Note that the trajectory of the pre-image of a smooth curve in Φ is a family of density functions on G, and the deformation of such functions may not necessarily be continuous. In particular, modes of the density may split and merge, appear and disappear.

4 Test study

Our goal in this section is to isolate a simple instance of the problem where our hypotheses can be verified. We therefore concentrate on a one-dimensional example, that is a simple geometric object that can be described by a one-dimensional parameter and mapped to a scalar label: $M = N = 1$. This allows us to easily visualize the maps from geometric to perceptual coordinates and their point-inverse.

In Figure 4 we show the test data that was presented to $K = 20$ subjects who were asked to rank the perceived assertiveness ϕ of the test subject on a scale from 1 to 10. Each subject performed 40 randomized tests, the first 10 of which were discarded from the final statistics since the subjects required a few samples to calibrate the scale. The data was obtained by

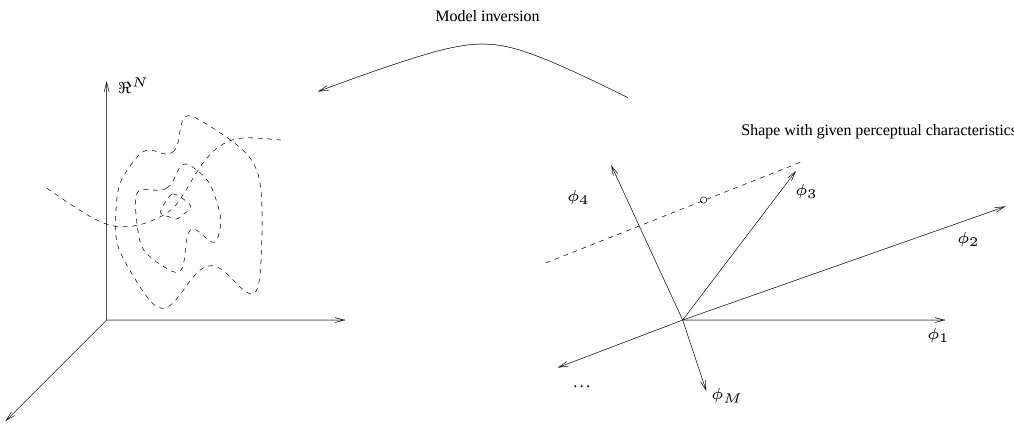

Figure 3: *A point in perceptual space induces a distribution in the corresponding shape: the levelsets of the density are shown as closed dashed curves. Changing perceptual characteristics induces changes in geometric ones. This is shown pictorially as the mode of the density describing a trajectory in geometric space as the pre-image of a straight line in perceptual space. Modes can split and merge and in general the pre-image can be arbitrarily complex.*

superimposing to a photograph of a female model downloaded from the Web a closed contour representing the rim of a set of eyeglasses. The upper edge of the rim consists of a Bezier curve with two fixed points (the hinges) and two variable tangents. The horizontal coordinate x of the intersection of the two tangents was randomized uniformly in an interval discretized into 10 segments. The histograms of the results of the experiments are shown in Figure 5. The histograms were then interpolated using a mixture of Gaussians. Conversely, for each point in perceptual coordinates ϕ, the pre-image of the experiment is shown in Figure 6. As it can be seen, the pre-image density is to a good approximation unimodal. Since the pre-image density is unimodal, we use its mode as the pointwise inverse map, and use the function as the map from perceptual to geometric coordinates. This map would allow a designer to vary the amount of perceived assertiveness without directly manipulating geometry.

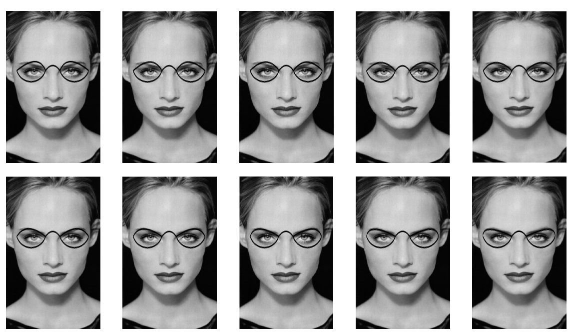

Figure 4: *Test data for the one-dimensional experiment "assertiveness". Randomize frames were presented to subjects in a forced-choice paradigm.*

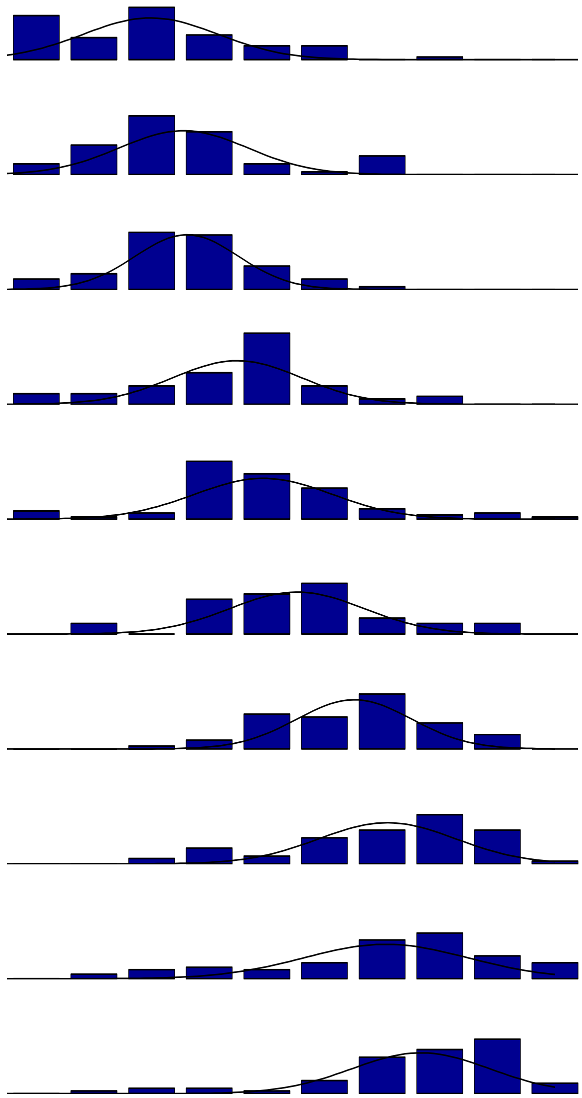

Figure 5: *Histograms of the map from geometric to perceptual coordinates.*

In the next experiment (Figure 7), we consider a more ambiguous "label": we have presented subjects with images of the same face where we randomized the height of the frames, and we asked them to quantify how "fashionable" the person wearing the frames looks. The histograms are shown in Figure 8 for 10 values of lens height. As it can be seen, the densities split into two modes at the extrema: this means that frames that are very narrow and very tall are either considered to be very fashionable or very unfashionable. This can be seen by following the modes in Figure 9, that bifurcates at the extrema.

5 Discussion

Due to the unusual nature of this work, we expect that several criticisms will be raised for which we have no satisfactory answer. For instance, is this a paper in Computer Vision/Pattern Recognition? Is the one-dimensional experiment discussed in the test case representative, or would the map inversion process described be subject to the curse of dimensionality when one wanted to integrate out nuisance parameters?

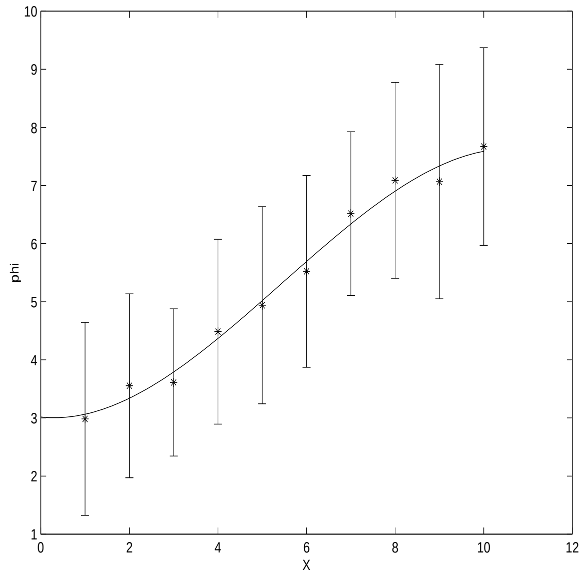

Figure 6: *Pre-image histograms mapping perceptual coordinates to geometry. Notice that the histograms are well approximated by unimodal densities.*

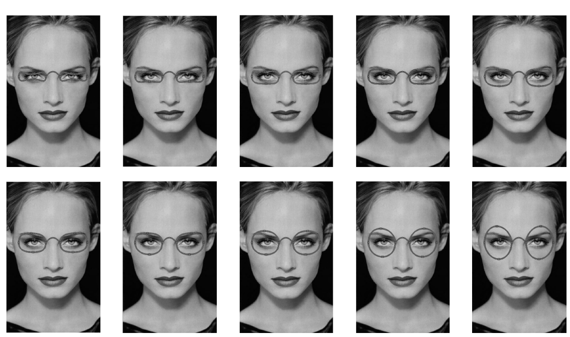

Figure 7: *Test data for the experiment “fashionable”. The images were randomized and presented to subjects in a forced-choice paradigm.*

One issue that we would like to stress is that the fact that the map recovered depends upon how, where, when and why the psychophysical data are collected is a desirable feature, and *not* a limitation. Our goal is to map a particular instance of an experiment, and we do not try to capture "universal" structure. For instance, in the case of eyeglasses, one cares to capture the taste and perception of a particular clientele, not of humanity at large.

We see the unsolved issues as opportunities for further research and hope that they will not detract from the attempt of quantifying and controlling the perceptual response to geometric features. The potential payoff in controlling the perceptual outcome in engineering design as well as other applications is, in our view, phenomenal, and we think that Computer Vision and Pattern Recognition ought to play a key role in the process.

Acknowledgements

We wish to thank Steven Engel for discussions and suggestions. This research is supported in part by NSF grant IIS-9876145, ARO grant DAAD19-99-1-0139 and Intel grant 8029.

References

[1] V. Blank and M. Vetter. In *Proc. of the ACM SIGGRAPH.*, 2000.

[2] M. Brand and L. Birnbaum. Perception as a matter of design. In *Working notes of the AAAI spring symposium on control of selective perception.*, pages 12–16, 1992.

[3] F. J. Gall. *Philosophisch-medicinische Untersuchungen uber Natur und Kunst im gesunden und kranken Zustande des Menschen.* 1 vol. Vienna, 1791. 2nd printing, Leipzig, 1800, 1791.

[4] F. J. Gall and J. G. Spurzheim. *Des dispositions innees de l'ame et de l'esprit : du materialisme, du fatalisme et de la liberte morale, avec des reflexions sur l'education et sur la llgislation criminelle*. Paris, 1811.

[5] J. Gibson. *The senses considered as perceptual systems*. Houghton Mifflin, Bosgon, 1966.

[6] W. Gogel. A theory of phenomenal geometry and its applications. *Perception and psychophysics*, 48:105–123, 1990.

[7] D. Goldberg, D. Nichols, B. Oki, and D. Terry. Using collaborative filering to weave an information tapestry. *C-ACM*, 35(12):61–70, 1992.

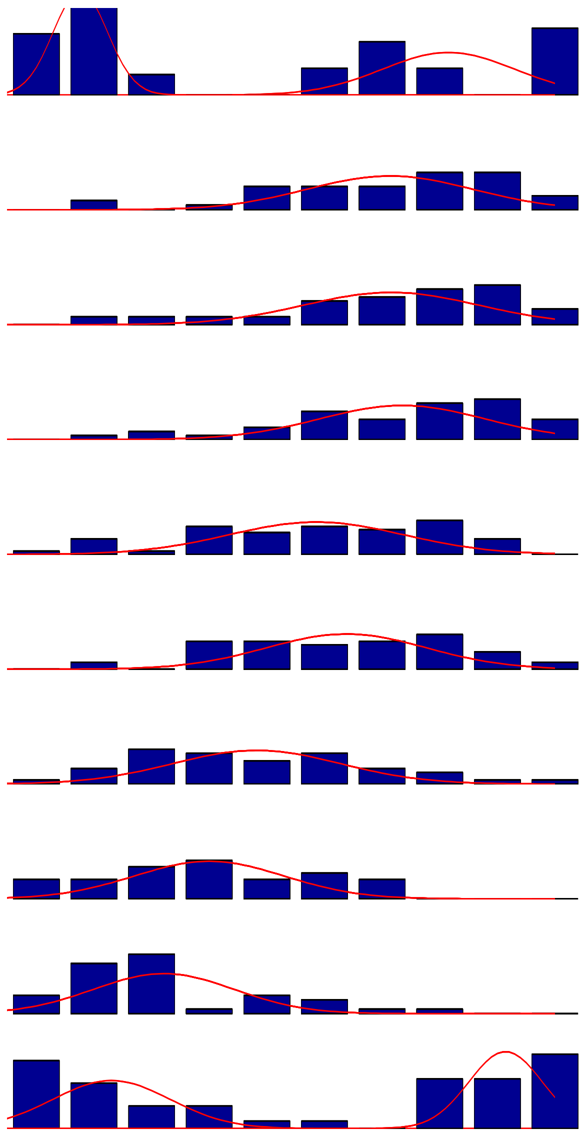

Figure 8: *Histograms for the experiment "fashionable": the modes bifurcates, indicating that excessively thin or tall glasses can be perceived equally likely as very fashionable or very unfashionable.*

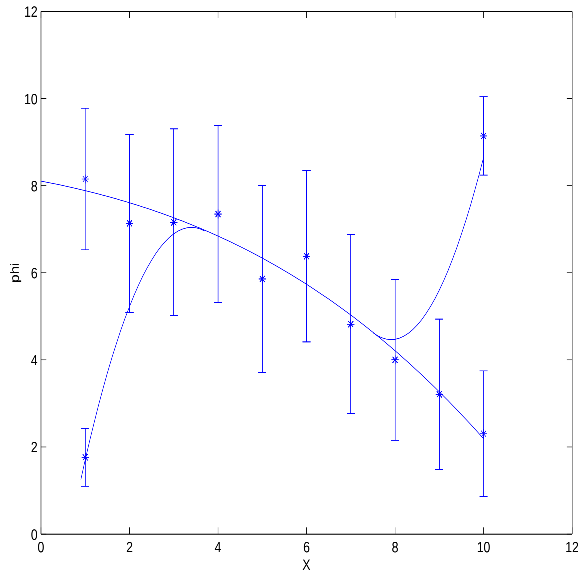

Figure 9: *Trajectory of the modes of the density for the experiment “fashionable”: the modes bifurcate.*